

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR LETTERS PATENT

**Method and Apparatus for Synchronizing  
Audio and Video Data**

Inventor(s):

James Lundblad

Ramaneek Khanna

ATTORNEY'S DOCKET NO. MS1-1051US

1 **TECHNICAL FIELD**

2 This invention relates to synchronizing audio data such that the audio data  
3 is played with the appropriate video data.  
4

5 **BACKGROUND OF THE INVENTION**

6 Various types of data streams contain both encoded video data and encoded  
7 audio data. Typically, a particular portion of the video data in a data stream  
8 corresponds with a particular portion of the audio data in the data stream. For  
9 example, if the video data is displaying a particular person speaking, the  
10 corresponding audio data presents the words or sounds uttered by that particular  
11 person. In this example, the presentation of the audio data should be synchronized  
12 with the presentation of the video data such that the movement of the speaker's  
13 lips at a particular moment corresponds to the word or sound being uttered.

14 A decoding device, such as a set-top box or other computing device,  
15 receives a data stream and decodes the video data and audio data contained in the  
16 data stream. The time required to decode and process the video data may differ  
17 from the time required to decode and process the audio data. This time difference  
18 may occur due to differences in the hardware components and/or software routines  
19 that process the video data and the audio data. Additionally, a particular time  
20 period of video data (e.g., one second) typically contains substantially more data  
21 than the same time period of audio data. Thus, the video data typically requires  
22 more processing than the audio data. Since the audio data may be processed faster  
23 than the video data, the audio data may not be ready for presentation while the  
24 video data is still being processed.  
25

1        Additionally, different clock signals (having different frequencies) may be  
2        used for processing the video data and the audio data. If these clocks are not  
3        synchronized, the audio data and video data may not be processed at the same rate,  
4        thereby adding to the uncertainty of the timing relationship between the video data  
5        and analog data.

6        Therefore it is desirable to provide a delay mechanism that adjusts the  
7        presentation of the audio data and/or the presentation of the video data such that  
8        the audio data is presented in synchronization with the appropriate video data.

## 9 10 11        **SUMMARY OF THE INVENTION**

12        The systems and methods described herein synchronize the presentation of  
13        audio data with the appropriate video data by determining a video presentation  
14        delay associated with the processing of the video data. The value of the video  
15        presentation delay is used to delay the presentation of the corresponding audio data  
16        such that the audio data is presented as substantially the same time as the  
17        associated video data.

18        In one embodiment, a transport stream is received containing video data and  
19        audio data. This embodiment determines the time required to process the video  
20        data contained in the transport stream and the time required to process the audio  
21        data contained in the transport stream. A determination is made regarding the  
22        difference in time to process the video data contained in the transport stream as  
23        compared to the audio data contained in the transport stream. Presentation of the  
24        audio data is delayed by the difference in time to process the video data contained  
25

1 in the transport stream as compared to the audio data contained in the transport  
2 stream.

3 According to one aspect of the invention, the determining of a time required  
4 to process the video data contained in the transport stream includes calculating a  
5 video presentation delay by comparing a presentation time stamp and a system  
6 time clock.

7 In a particular embodiment, delaying presentation of the audio data includes  
8 storing the audio data in a buffer with a delay that corresponds to the difference in  
9 time to process the video data contained in the transport stream as compared to the  
10 audio data contained in the transport stream.

### 11 **BRIEF DESCRIPTION OF THE DRAWINGS**

12  
13 The same reference numerals are used throughout the drawings to reference  
14 like components and features.

15  
16 Fig. 1 illustrates an exemplary environment in which the methods and  
17 systems described herein may be implemented.

18 Fig. 2 is a block diagram of an example client device, a television, and  
19 various input devices that interact with the client device.

20 Fig. 3 is a block diagram of selected components of the client device shown  
21 in Figs. 1 and 2.

22 Fig. 4 is a block diagram of an exemplary system that decodes transport  
23 streams.

24 Fig. 5 is a flow diagram illustrating an embodiment of a procedure for  
25 synchronizing an audio signal with a video signal.

1 Fig. 6 is a block diagram of an exemplary system for processing a video  
2 portion of a transport stream.

3 Fig. 7 is a flow diagram illustrating an embodiment of a procedure for  
4 processing a video portion of a transport stream using the system shown in Fig. 6.

5 Fig. 8 is a block diagram of an exemplary system for processing an audio  
6 portion of a transport stream.

7 Fig. 9 is a flow diagram illustrating an embodiment of a procedure for  
8 processing an audio portion of a transport stream using the system shown in Fig. 8.

## 9 10 11 **DETAILED DESCRIPTION**

12 Fig. 1 illustrates an exemplary environment 100 in which the methods and  
13 systems described herein may be implemented. One or more content providers 102  
14 include stored content 118 and a content server 120. Content server 120 controls  
15 the movement of content (including stored content 118) from the content provider  
16 102 to a content distribution system 104, which is coupled to the content provider.  
17 Additionally, the content server 120 controls the movement of live content (e.g.,  
18 content that was not previously stored by the content provider) and content stored  
19 at other locations to the content distribution system.

20 The content distribution system 104 contains a broadcast transmitter 122  
21 and one or more content processors 124. Broadcast transmitter 122 broadcasts  
22 signals (e.g., cable television signals) across a broadcast network 116, such as a  
23 cable television network. Broadcast network 116 may include wired or wireless  
24 media using any broadcast format or broadcast protocol. Content processor 124  
25 processes the content received from content provider 102 prior to transmitting the

1 content across the broadcast network 116. A particular content processor may  
2 encode or otherwise process the received content into a format that is understood  
3 by multiple client devices 106 coupled to the broadcast network 116. Although  
4 Fig. 1 shows a single content provider 102 and a single content distribution system  
5 104, a particular environment may include any number of content providers  
6 coupled to any number of content distribution systems.

7 A client device 106(1) receives broadcast content from a satellite-based  
8 transmitter via a satellite dish 110. Client device 106(1) is also referred to as a set-  
9 top box, game console or a satellite receiving device. Client device 106(1) is  
10 coupled to a television 108(1) for presenting the content received by the client  
11 device (i.e., audio data and video data) as well as a graphical user interface. A  
12 particular client device 106 may be coupled to any number of televisions 108.  
13 Similarly, any number of client devices 106 may be coupled to a television 108.  
14 Another client device 106(2) is coupled to receive broadcast content from  
15 broadcast network 116 and provide the received content to a television 108(2).  
16 Another client device 106(N) is a combination of a television 112 and a set-top  
17 box 114. In this example, the various components and functionality of the set-top  
18 box are incorporated into the television, rather than using two separate devices.  
19 The set-top box incorporated into the television may receive broadcast signals via  
20 a satellite dish (similar to satellite dish 110) and/or via broadcast network 116. In  
21 alternate embodiments, client devices 106 may receive broadcast signals via the  
22 Internet or any other broadcast medium.

23 Fig. 2 is a block diagram of an example client device 106, television 108,  
24 and various input devices that interact with the client device. As discussed above,  
25 client device 106 may also be referred to as a set-top box, game console or a

1 satellite receiver. Client device 106 includes a wireless receiving port 202 (e.g., an  
2 infrared (IR) wireless port) for receiving wireless communications from a remote  
3 control device 204, a handheld device 206 (such as a personal digital assistant  
4 (PDA) or handheld computer), or other wireless device, such as a wireless  
5 keyboard. Additionally, a wired keyboard 208 is coupled to client device 106 for  
6 communicating with the client device. In alternate embodiments, remote control  
7 device 204, handheld device 206, and/or keyboard 208 may use an RF  
8 communication link (or other mode of transmission) to communicate with client  
9 device 106.

10 Client device 106 receives one or more broadcast signals 220 from one or  
11 more broadcast sources (e.g., from a broadcast network or via satellite). Client  
12 device 106 includes hardware and/or software for receiving and decoding  
13 broadcast signal 220, such as an NTSC, PAL, SECAM or other TV system video  
14 signal, and providing video data to the television 108. Client device 106 also  
15 includes hardware and/or software for providing the user with a graphical user  
16 interface by which the user can, for example, access various network services,  
17 configure the client device 106, and perform other functions.

18 Client device 106 receives AC power on line 110. Client device 106 is  
19 capable of communicating with other devices via a conventional telephone link  
20 212, an ISDN link 214, a cable link 216, and an Ethernet link 218. A particular  
21 client device 106 may use any one or more of the various communication links  
22 212-218 at a particular instant. Client device 106 also generates a video signal and  
23 an audio signal, both of which are communicated to television 108. The video  
24 signals and audio signals can be communicated from client device 106 to  
25 television 108 via an RF (radio frequency) link, S-video link, composite video link,

1 component video link, or other communication link. Although not shown in Fig.  
2 2, a particular client device 106 may include one or more lights or other indicators  
3 identifying the current status of the client device. Additionally, a particular client  
4 device 106 may include one or more control buttons or switches (not shown) for  
5 controlling operation of the client device.

6 Fig. 3 is a block diagram of selected components of the client device 106  
7 shown in Figs. 1 and 2. Client device 106 includes a first tuner 300 and an  
8 optional second tuner 302, one or more processors 304, a random access memory  
9 (RAM) 306, and a non-volatile memory 308 that contains, for example, an  
10 operating system 310 and one or more application programs 312. Client device  
11 106 also includes a disk drive 314 and storage media 316. Although client device  
12 106 is illustrated having both a RAM 306 and a disk drive 314, a particular device  
13 may include only one of the memory components. Additionally, although not  
14 shown, a system bus typically couples together the various components within  
15 client device 106.

16 Processor(s) 304 process various instructions to control the operation of  
17 client device 106 and to communicate with other electronic and computing  
18 devices. The memory components (e.g., RAM 306, disk drive 314, storage media  
19 316, and non-volatile memory 308) store various information and/or data such as  
20 configuration information and graphical user interface information.

21 Client device 106 also includes a decoder 318, such as an MPEG-2 decoder  
22 that decodes MPEG-2-encoded signals. A modem 320 allows client device 106 to  
23 communicate with other devices via a conventional telephone line. An IR  
24 interface 322 allows client device 106 to receive input commands and other  
25 information from a user-operated device, such as a remote control device or an IR



1 keyboard. Client device 106 also includes a network interface 324, a serial/parallel  
2 interface 326, an audio output 328, and a video output 330. Interfaces 324 and 326  
3 allow the client device 106 to interact with other devices via various  
4 communication links. Although not shown, client device 106 may also include  
5 other types of data communication interfaces to interact with other devices. Audio  
6 output 328 and video output 330 provide signals to a television or other device that  
7 processes and/or presents the audio and video data. Although client 106 is  
8 illustrated having multiple interfaces, a particular client may only include one or  
9 two such interfaces.

10 Client device 106 also includes a user interface (not shown) that allows a  
11 user to interact with the client device. The user interface may include indicators  
12 and/or a series of buttons, switches, or other selectable controls that are  
13 manipulated by a user of the client device.

14 General reference is made herein to one or more client devices, such as  
15 client device 106. As used herein, "client device" means any electronic device  
16 having data communications, data storage capabilities, and/or functions to process  
17 signals, such as broadcast signals, received from any of a number of different  
18 sources.

19 Fig. 4 is a block diagram of an exemplary system 400 that decodes one or  
20 more transport streams. A "transport stream" may also be referred to as a  
21 "program stream" or a "data stream". System 400 may use one or more of the  
22 components shown in Fig. 3, such as processor(s) 304, application program(s) 312,  
23 and decoder 318. A transport stream decoder 402 receives a transport stream, such  
24 as an MPEG-2 data stream, and separates the video and audio portions of the  
25 transport stream. Transport stream decoder 402 provides the video portion of the

1 transport stream to a video processing module 406 and provides the audio portion  
2 of the transport stream to an audio processing module 408. Video processing  
3 module 406 handles the decoding of the video portion of the transport stream and  
4 generates decoded video data that is formatted for display on a display device,  
5 such as a television. Audio processing module 408 handles the decoding of the  
6 audio portion of the transport stream and generates decoded audio data that is  
7 formatted for broadcast by a broadcast device, such as one or more speakers in a  
8 television.

9 The transport stream also includes timing information (e.g., time stamps)  
10 that is extracted by transport stream decoder 402 and provided to a clock control  
11 module 404. Clock control module 404 provides one or more control signals to  
12 video processing module 406 and audio processing module 408 to synchronize the  
13 decoded video data with the decoded audio data.

14 A particular embodiment of the invention will be described in the context of  
15 a transport stream encoded using the MPEG-2 (Moving Pictures Experts Group).  
16 MPEG-2 is a standard for digital video and digital audio compression. MPEG-2  
17 supports a variety of audio/video formats, including legacy TV, HDTV (High-  
18 Definition Television), and five channel surround sound. For example, MPEG-2 is  
19 capable of providing broadcast-quality images of 720x480 resolution used in DVD  
20 movies. However, the methods and systems described herein can be used with any  
21 type of data stream using any type of encoding format as well as data streams that  
22 do not use any encoding.

23 A particular broadcast format provides for the transmission of X image  
24 frames per second, such as 30 frames per second or 60 frames per second. A  
25

1 particular frame includes two interlaced fields, in which each field includes a  
2 specific number of horizontal scan lines. The broadcast and display of image  
3 frames is described in connection with a conventional analog television having a  
4 cathode ray tube (CRT) with an electron beam. The electron beam is controlled  
5 such that the electron beam is scanned across the screen of the CRT to generate the  
6 appropriate image.

7 The first few horizontal scan lines may be used to synchronize the  
8 television receiver and to return the electron beam to the top of the screen. The  
9 electron beam is disabled (also referred to as “blanked”) during this time so that  
10 the electron beam does not generate a visible line from the bottom of the screen to  
11 the top of the screen when being returned to the top of the screen. These first few  
12 horizontal scan lines are commonly referred to as the “vertical blanking interval”  
13 lines (or VBI lines).

14 The odd scan lines of the frame (i.e., frame line 1, frame line 3, etc.) are  
15 received first and are referred to as the “odd field”. A particular number of these  
16 odd lines are the VBI lines. The VBI lines synchronize the television receiver for  
17 the subsequent scanning of the horizontal scan lines of a viewable portion of the  
18 frame. For each horizontal scan line, the electron beam scans from left to right  
19 across the screen. When the electron beam reaches the right edge of the screen,  
20 the electron beam is returned to the left edge of the screen in preparation for the  
21 scanning of the next scan line. After the scanning of each odd scan line in the  
22 viewable portion, the electron beam is “blanked” as the electron beam is returned  
23 to left edge of the screen in preparation for the start of the next scan line. This  
24 blanking time is referred to as the “horizontal blanking interval” of the frame.  
25

1 After the last odd scan line has finished, the even scan lines of the frame  
2 (i.e., frame line 2, frame line 4, etc.) are received and are referred to as the "even  
3 field". As with the odd field discussed above, a particular number of the scan lines  
4 of the even field are VBI lines. The electron beam is blanked during the scanning  
5 of the even VBI lines such that the electron beam can be returned to the top of the  
6 screen without generating a line on the screen. After the scanning of all the even  
7 VBI lines, the even scan lines of the viewable portion are scanned in a manner  
8 similar to the scanning of the odd scan lines discussed above. The viewable  
9 horizontal scan lines of the odd and even fields together cause the electron beam to  
10 scan across the screen of the television to create the viewable television image.  
11 Although the example described above applies to interlaced video signals, the  
12 methods and systems described herein can be used with both interlaced and non-  
13 interlaced video signals.

14 Referring again to Fig. 4, there is a video processing delay that is defined as  
15 the time required to process (using hardware and/or software) the video portion of  
16 a received transport stream. With reference to Fig. 4, the video processing delay is  
17 the time that elapses between receiving a particular set of video data at the  
18 transport stream decoder 402 and outputting the corresponding decoded video data  
19 from the video processing module 406. Similarly, there is an audio processing  
20 delay that is defined as the time required to process (using hardware and/or  
21 software) the audio portion of a received transport stream. With reference to Fig.  
22 4, the audio processing delay is the time that elapses between receiving a particular  
23 set of audio data at the transport stream decoder 402 and outputting the  
24 corresponding decoded audio data from the audio processing module 408. The  
25

1 video processing delay and the audio processing delay may include decoder  
2 buffering delays, decoding delays, and/or presentation delays.

3 Fig. 5 is a flow diagram illustrating an embodiment of a procedure 500 for  
4 synchronizing an audio signal with a video signal. Initially, procedure 500  
5 receives a transport stream containing encoded video data and encoded audio data  
6 (block 502). The transport stream may be received, for example, via a broadcast  
7 network, such as a cable television network, or via a satellite transmission system.  
8 The procedure 500 determines the time required to process the video portion of the  
9 transport stream (block 504). Next, the procedure determines the time required to  
10 process the audio portion of the transport stream (block 506). The procedure then  
11 determines the difference in time to process the video portion of the transport  
12 stream as compared to the audio portion of the transport stream (block 508).  
13 Block 510 then determines which processing time is greater (i.e., the video  
14 processing time determined at block 504 or the audio processing time determined  
15 at block 506). If the audio processing time is greater, the video presentation is  
16 delayed by the difference determined at block 508, thereby synchronizing the  
17 decoded video data with the decoded audio data. If the video processing time is  
18 greater, the audio presentation is delayed by the difference determined at block  
19 508, thereby synchronizing the decoded audio data with the decoded video data.  
20 Additional details regarding the various actions described above with respect to  
21 Fig. 5 are provided below with reference to Figs. 6-9.

22 In a particular embodiment, the decoded audio data is “substantially  
23 synchronized” with the decoded video data. “Substantially synchronized” means  
24 that there may be a slight difference (such as a few milliseconds) between the  
25 presentation of the video data and the presentation of the corresponding audio data.

1 Such a small difference in the presentation of the audio and video data is not likely  
2 to be perceived by a user watching and listening to the presented video and audio  
3 data.

4 A typical transport stream is received at a substantially constant rate. In this  
5 situation, the delay that is applied to the video presentation or the audio  
6 presentation is not likely to change frequently. Thus, the procedure of Fig. 5 may  
7 be performed periodically (e.g., every few seconds or every 30 received video  
8 frames) to be sure that the delay currently being applied to the video presentation  
9 or the audio presentation is still within a particular threshold (e.g., within a few  
10 milliseconds of the required delay). Alternatively, the procedure of Fig. 5 may be  
11 performed for each new frame of video data received from the transport stream.

12 In another embodiment, the procedure of Fig. 5 is performed as described  
13 above, but the audio or video presentation delay is changed only if the newly  
14 calculated delay value exceeds the delay value currently being used by a threshold  
15 value (e.g., ten milliseconds). Thus, although the delay is recalculated frequently,  
16 the actual delay applied by the system is only changed when the new delay exceeds  
17 the value.

18 Typically, video data processing requires more time than audio data  
19 processing. Thus, in an alternative embodiment where the video processing time is  
20 known to be greater than the audio processing time, blocks 510 and 512 of Fig. 5  
21 can be eliminated. In this embodiment, the difference determined in 508 is used to  
22 determine an additional delay that is applied to the audio presentation. Without  
23 this additional delay, the audio data might be presented to the user prior to the  
24 associated video data (i.e., not synchronized).

1 In a typical MPEG-2 transport stream, the timing is defined in terms of a  
2 common system clock, referred to as a System Time Clock (STC).  
3 Synchronization of audio and video data is accomplished using Presentation Time  
4 Stamps (PTS) contained in the transport stream. In a particular embodiment, an  
5 MPEG-2 transport stream has an associated system clock frequency of 27 MHz ( $\pm$   
6 810 Hz). Thus, a bit rate of 27,000,000 bits per second indicates that one byte of  
7 data is transferred every eight cycles of the system clock.

8 Fig. 6 is a block diagram of an exemplary system 600 for processing a video  
9 portion of a transport stream. A video clock module 602 receives a reference time  
10 stamp (RTS), which is contained in the MPEG-2 transport stream. The video  
11 clock module 602 is locked to the RTS in the transport stream. Video clock  
12 module 602 generates a timing reference signal that is provided to a video timing  
13 generator 604 and video display hardware 606. Video timing generator 604  
14 generates one or more sync signals used by the video display hardware 606 to  
15 format the video output to the television. Video timing generator 604 also  
16 generates a VSYNC (vertical retrace sync) signal, which generates a software  
17 interrupt used by a video display software routine 608. The VSYNC signal is  
18 generated each time a complete image field (e.g., an odd field or an even field) has  
19 been rendered and the electron beam is returned to the beginning of the CRT to  
20 begin rendering the next image field. Alternatively, the VSYNC signal may be  
21 generated each time a complete frame has been rendered.

22 The video display hardware 606 receives the video portion of the transport  
23 stream (e.g., by reading the received video frame from a video memory device).  
24 The video portion of the transport stream represents decoded video data. The  
25 video decoding can be performed in hardware, software, or a combination of

1 hardware and software. In a particular embodiment, the video decoding is  
2 performed by the transport stream decoder 402 (Fig. 4).

3 Video display hardware 606 also receives information from video display  
4 software routine 608 regarding when to display the next frame of video data. The  
5 video data is formatted and converted to an analog video signal that is  
6 synchronized to the video timing generator 604. The analog video signal is output  
7 from the video display hardware 606 to a television or other display device.

8 The video display software routine 608 receives the VSYNC signal from  
9 the video timing generator 604. When the VSYNC interrupt occurs, a time stamp  
10 is taken from a CPU clock 612. The CPU clock is a free running clock based on  
11 the CPU bus frequency. The CPU clock can be read, for example, via a kernel  
12 API. The time stamp resulting from the VSYNC interrupt is used as a reference  
13 for a system time clock (STC) 610. The system time clock (STC) is derived from  
14 the video timing generator 604 (using the VSYNC interrupt) and the CPU clock  
15 612. For each VSYNC interrupt, the STC is advanced the number of ticks in one  
16 field time (i.e., the number of clock cycles required to transmit a full field of data  
17 in the transport stream). The CPU clock is used to interpolate the appropriate  
18 number of ticks between VSYNC interrupts. Since the frequency of the MPEG  
19 data transmission frequency is known (27 MHz), and the amount of data bytes  
20 required to fill a field of data is known, the number of ticks to advance the STC  
21 can be determined. The formula to calculate the number of ticks to advance the  
22 STC clock is as follows:

23  
24 
$$\text{No. of Ticks to Advance} = T_{\text{field}} * 27,000,000$$
  
25



1 In the United State,  $T_{\text{field}} = 16.6833333\ldots$  milliseconds.

2 The video display software routine 608 compares the presentation time  
3 stamp (PTS) encoded in the video frame and the system time clock 610 at the time  
4 of the VSYNC interrupt. The difference in time between the PTS and the STC at  
5 the time of the VSYNC interrupt is the video presentation delay, which is provided  
6 to the audio processing system to delay the audio output by the video presentation  
7 delay, thereby synchronizing the audio output with the video output .

8 Fig. 7 is a flow diagram illustrating an embodiment of a procedure 700 for  
9 processing a video portion of a transport stream using the system shown in Fig. 6.  
10 Initially, the procedure receives reference time stamps (RTS) from a transport  
11 stream (block 702). The procedure then generates synchronization signals used to  
12 format the video data from the transport stream for output to a television or other  
13 display device (block 704). The procedure generates a software interrupt each  
14 time a VSYNC signal is received (block 706). At block 708, the procedure  
15 provides a next frame of video data to the video display hardware for processing.  
16 This processing by the video display hardware may be performed concurrently  
17 with the remaining activities of procedure 700.

18 The procedure then determines whether a software interrupt has been  
19 received (block 710). If not, the procedure awaits the next software interrupt. If a  
20 software interrupt has been received, the procedure retrieves a time stamp from a  
21 CPU clock (block 712). A presentation time stamp (PTS) is compared with the  
22 CPU clock time stamp (block 714). A video presentation delay is generated that  
23 represents the difference between the PTS and the CPU clock time stamp (block  
24 716).

Fig. 8 is a block diagram of an exemplary system 800 for processing an audio portion of a transport stream. An audio clock module 802 is locked to the reference time stamp (RTS) contained in the transport stream. The audio clock module 802 generates a timing reference used by audio reproduction hardware 804, along with other data, to generate an analog audio signal that is provided to, for example, a television. The audio reproduction hardware 804 receives audio data from one or more DMA buffers 812, which are controlled by a DMA controller 810. The audio reproduction hardware 804 converts the data received from DMA buffers 812 into an analog audio signal.

An audio software routine 806 is coupled to the DMA controller 810 and a system time clock 610 (e.g., the same system time clock shown in Fig. 6). Audio software routine 806 receives presentation time stamps (PTS) from the transport stream and receives video presentation delay information generated by the video display software routine 608 shown in Fig. 6. Audio software routine 806 controls the placement of decoded audio frames in the DMA buffers 812 (via DMA controller 810) with a delay matching the video presentation delay reported by the video display software routine. Specifically, audio software routine 806 reads a presentation time stamp from each audio frame before it is decoded. The audio software routine 806 then reads the system time clock 610, the video presentation delay, and the position of the DMA read pointer (provided by the DMA controller 810). The audio frame is then decoded and stored in the DMA buffers 812 with a delay that matches the video presentation delay. The audio data is decoded in, for example, audio software routine 806. Alternatively, the audio data may be decoded in hardware or a combination of hardware and software.

1 Fig. 9 is a flow diagram illustrating an embodiment of a procedure 900 for  
2 processing an audio portion of a transport stream using the system shown in Fig. 8.  
3 Initially, procedure 900 receives reference time stamps (RTS) from a transport  
4 stream (block 902). The procedure then generates timing signals used to generate  
5 an analog audio signal (block 904). Presentation time stamps (PTS) are then  
6 received from the transport stream (block 906). The procedure also receives video  
7 presentation delay information generated by the video display software routine  
8 (block 908).

9 The procedure 900 then decodes the audio data contained in the transport  
10 stream (block 910). The decoded audio data is then stored in one or more DMA  
11 buffers with a delay matching the video presentation delay (block 912). At the  
12 appropriate time, the audio data is provided from the DMA buffers to the audio  
13 reproduction hardware (block 914). The audio reproduction hardware converts the  
14 audio data to an analog signal that can be provided to a presentation device, such  
15 as the speakers in a television.

16 Portions of the systems and methods described herein may be implemented  
17 in hardware or a combination of hardware, software, and/or firmware. For  
18 example, one or more application specific integrated circuits (ASICs) or  
19 programmable logic devices (PLDs) could be designed or programmed to  
20 implement one or more portions of the video and/or audio processing systems and  
21 procedures.

22 Although the invention has been described in language specific to structural  
23 features and/or methodological steps, it is to be understood that the invention  
24 defined in the appended claims is not necessarily limited to the specific features or  
25

1 steps described. Rather, the specific features and steps are disclosed as preferred  
2 forms of implementing the claimed invention.

3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25